REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
	Reprint	

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Following Navigation Instructions Presented Verbally or Spatially: Effects on Training, Retention and Transfer	W911NF-05-1-0153
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER 611103
6. AUTHORS Vivian I. Schneider, Alice F. Healy, Immanuel Barshi, James A. Kole	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES	8. PERFORMING ORGANIZATION REPORT NUMBER
University of Colorado - Boulder Office of Contracts and Grants Campus Box 572, 3100 Marine Street Rm 481 Boulder, CO 80309 -0572	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) ARO
U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211	11. SPONSOR/MONITOR'S REPORT NUMBER(S) 48097-MA-MUR.50

12. DISTRIBUTION AVAILIBILITY STATEMENT

Approved for Public Release;federal purpose rights

13. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT

Two experiments investigated participants' ability to follow navigation instructions in a situation simulating communication between air traffic controllers and aircrews. A verbal condition, in which instructions were given orally, was compared with a spatial condition, in which commands were shown on a computer display as simulated movements, with the presentation times in the two conditions equated. Retention and transfer were studied a week later when participants performed in either the same or the other condition. In both sessions, participants' initial

15. SUBJECT TERMS

navigation instructions, skill training, skill retention, skill transfer, verbal presentation, spatial presentation

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Alice Healy
UU	UU	UU	UU		19b. TELEPHONE NUMBER 303-492-5032

Standard Form 298 (Rev 8/98)
Prescribed by ANSI Std. Z39.18

Report Title

Following Navigation Instructions Presented Verbally or Spatially: Effects on Training, Retention and Transfer

ABSTRACT

Two experiments investigated participants' ability to follow navigation instructions in a situation simulating communication between air traffic controllers and aircrews. A verbal condition, in which instructions were given orally, was compared with a spatial condition, in which commands were shown on a computer display as simulated movements, with the presentation times in the two conditions equated. Retention and transfer were studied a week later when participants performed in either the same or the other condition. In both sessions, participants' initial proportion correct was much higher in the spatial than in the verbal condition, but after three blocks, accuracy in the two conditions was equivalent. Retention was perfect when training and test conditions matched. Training in the verbal condition transferred to the spatial condition but not vice versa. Thus, there is evidence that participants' representations of the movements in the verbal and spatial conditions were not equivalent.

APPLIED COGNITIVE PSYCHOLOGY
Appl. Cognit. Psychol. (2009)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/acp.1642

Following Navigation Instructions Presented Verbally or Spatially: Effects on Training, Retention and Transfer[†]

VIVIAN I. SCHNEIDER[1]*, ALICE F. HEALY[1],
IMMANUEL BARSHI[2] and JAMES A. KOLE[1]

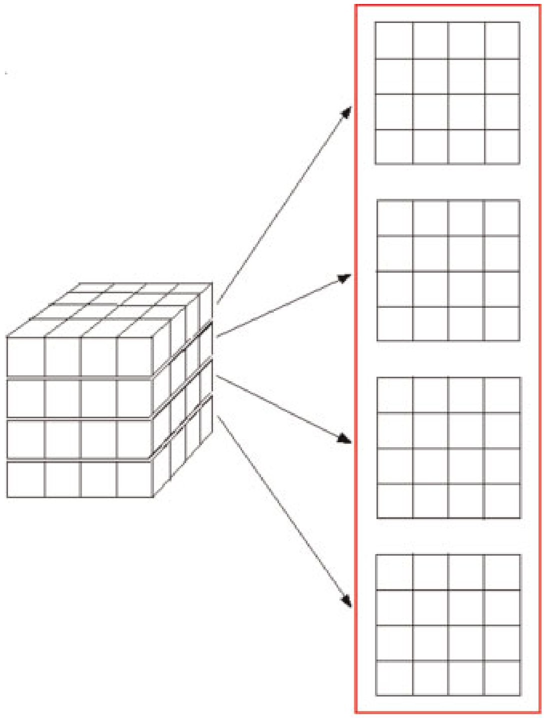
[1]*University of Colorado, Boulder, USA*
[2]*NASA, Ames Research Center, USA*

SUMMARY

Two experiments investigated participants' ability to follow navigation instructions in a situation simulating communication between air traffic controllers and aircrews. A verbal condition, in which instructions were given orally, was compared with a spatial condition, in which commands were shown on a computer display as simulated movements, with the presentation times in the two conditions equated. Retention and transfer were studied a week later when participants performed in either the same or the other condition. In both sessions, participants' initial proportion correct was much higher in the spatial than in the verbal condition, but after three blocks, accuracy in the two conditions was equivalent. Retention was perfect when training and test conditions matched. Training in the verbal condition transferred to the spatial condition but not *vice versa*. Thus, there is evidence that participants' representations of the movements in the verbal and spatial conditions were not equivalent. Published in 2009 by John Wiley & Sons, Ltd.

Communication between air traffic controllers and flight crews primarily involves giving and receiving navigation instructions. For example, air traffic control might dispatch in one message to the flight crew information about compass heading, speed, altitude restriction and approach clearance. The amount of information to be processed might exceed normal human capacity (Cowan, 2001), and this information must be held in memory for an extended temporal duration. Hence, flight crews sometimes make errors under these circumstances, which can lead to serious accidents. We have been studying this communication situation with the eventual goal to determine ways to reduce such critical errors. In particular, we have been investigating factors influencing participants' ability to follow navigation instructions in a paradigm meant to simulate communication between air traffic controllers and flight crews (see, e.g. Barshi & Healy, 1998, 2002; Healy, Schneider, & Barshi, 2009; Schneider, Healy, & Barshi, 2004). In the standard version of this paradigm, participants hear navigation instructions like those given by air traffic control about moving in a space shown on a computer monitor; they repeat aloud (or 'readback') the instructions, as pilots are expected to do; and then they follow the instructions, navigating in the space displayed on the computer. The instructions describe movements in a grid of four 4×4 matrices shown in Figure 1 (top panel). Enclosed in the box on the right is the display that the participant

*Correspondence to: Dr Vivian I. Schneider, Department of Psychology and Neuroscience, 345 UCB, University of Colorado, Boulder, CO 80309-0345, USA. E-mail: vivian.schneider@colorado.edu



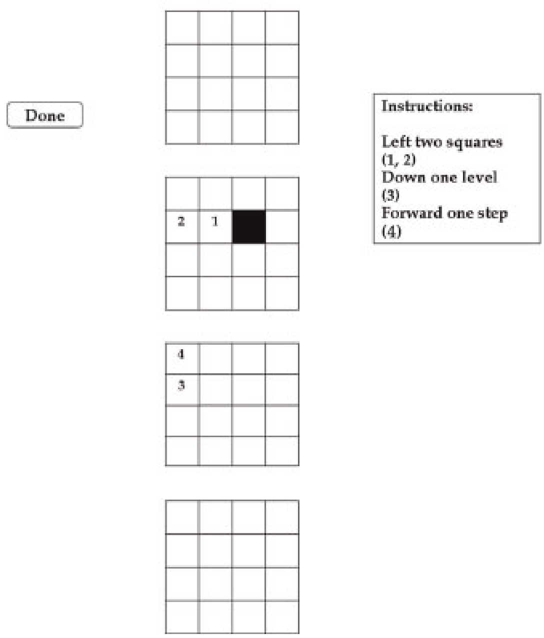


Figure 1. Sample screen display showing four stacked matrices (in the box on the right) and the three-dimensional space represented by the matrices (on the left) (top panel). Sample practice display with sample set of instructions including three commands. The participants did not see the instructions. The numerals shown in the instructions and on the matrices indicate the locations of the required clicks. The starting point is indicated by the filled-in square (bottom panel)

 Appl. Cognit. Psychol. (2009)
DOI: 10.1002/acp

sees in the experiment; also shown on the left is the three-dimensional space it represents. Participants hear instructions like the following one including three commands: Left two squares, down one level, forward one step. Upon hearing such instructions, the participant immediately repeats them, and, next, uses the computer mouse to follow the instructions, by clicking each appropriate square on the grid in the order specified (see Figure 1 bottom panel).

In studies using this paradigm, we have explored the relationship between the mental representations of both the words in the instructions and the movements required by them to navigate in the space. We found that the verbal representation of the instructions was influenced by the complexity of the spatial representation used to make movements. In particular when the movements required the participants to represent the space as a three-dimensional structure, as opposed to a simpler two-dimensional structure, the participants showed worse memory for the verbal instructions, as reflected in their readback responses (e.g. Barshi & Healy, 2002), even though the readback immediately followed hearing the instructions and preceded the movements through the space. Furthermore, when the starting point of the movements was known in advance of hearing the verbal instructions, participants showed better memory for those instructions, as reflected in their readback responses, than when the starting point was known only after the verbal instructions were given and repeated (Schneider, Healy, Kole, & Barshi, 2003), again even though the readback responses preceded the movements in the space.

The present study explores the nature of the relationship between participants' verbal and spatial representations of navigation instructions in a new way. We compared a verbal condition, in which navigation instructions were given orally, to a spatial condition, in which no verbal commands were given but participants made the same movements. In the spatial condition, the commands were shown visually on the computer as simulated movements. The instruction time in the two conditions was equated. In both conditions, the participants were to follow the commands or reproduce the movement path by using a mouse to click on each appropriate square on the grid in the order specified. Differences between verbally presenting movement instructions and presenting the same instructions in a non-verbal format have previously been explored in other paradigms (e.g. Farrell, 1979; Loftus, 1978; Maki, 1979; Mou, Zhang, & McNamara, 2004). We might expect that directly presenting the movement sequences spatially would yield the same strategies and, thus, comparable performance to that found when describing those sequences verbally so that instructions in the two modalities would show the same pattern of results. Alternatively, if non-negligible effort is required to create the spatial representation from the verbal representation (i.e. to translate from a verbal representation to the required spatial representation), then directly presenting the movement sequences spatially should yield better performance than when the sequences are described verbally, requiring that the spatial representation be created. This second possibility is also consistent with the stimulus/central-processing/response (S-C-R) compatibility model of a pilot's task based on the assumption that spatial tasks are more compatible with visual inputs than with auditory inputs (Wickens, Vidulich, & Sandry-Garza, 1984). A third possibility is that performance would be better with a verbal presentation of the required movements than with a presentation of the movement sequences themselves because the verbal presentation would involve advantages from dual coding of both verbal and spatial representations, whereas only a single, spatial code might be involved when the movements are presented directly (see, e.g. Paivio, 1991, 2007).

PRESENT STUDY

In Experiment 1, we compared a verbal condition, in which navigation instructions were given with verbal commands, to a spatial condition, in which no verbal commands were given but participants made the same movements. In the previously reported experiments using this paradigm (e.g. Barshi & Healy, 1998, 2002; Healy et al., 2009; Schneider et al., 2004), messages were varied in length from 1 to 6 commands (a command is a basic unit of movement that consists of a dimension, a direction and a magnitude of movement). It was found that performance decreased with increases in message length, with the largest decreases in the intermediate ranges of message length (i.e. 3–5 commands). Also, performance generally increased across blocks of trials, suggesting that training led participants to improve their skill for understanding, remembering and following navigation instructions in this paradigm.

In the previous studies, movement dimension was confounded with both message length and serial position in the sequence of commands. The commands always occurred in a fixed order with respect to the dimension of movement. Movements were always made first to the right or left, then up or down, then forward or back, then, repeated in the same order for longer message lengths. To examine the effects of movement dimension, in the present study we removed the previous confounding by allowing each movement dimension to occur equally often at all six message lengths and at all serial positions.

Previous studies of spatial information processing using other paradigms found the right/left dimension more difficult than the other dimensions for verbal stimuli but not for non-verbal stimuli (see, e.g. Loftus, 1978; Maki, 1979; Mou et al., 2004; but see Farrell, 1979). Likewise, Sholl and Egeth (1981) performed a series of experiments using orthogonal manipulations of verbal labels and spatial axes to disentangle verbal from perceptual encoding explanations for right/left difficulty, and they concluded that the difficulty is in the verbal encoding. If these results generalize to the present paradigm, we would expect to find that participants are less accurate on the right/left commands than on the commands along the other dimensions in the verbal condition but not in the spatial condition.

Errors in which participants confuse right and left are not the only type they can make. In our navigation paradigm, there were three kinds of movements participants could make: Lateral (right/left), vertical from one matrix to another (up/down) and (in terms of the display shown on the screen though not in terms of the three-dimensional space itself) vertical within a single matrix (back/forward). Thus, another possible error type is one in which participants move along the wrong dimension (e.g. they move up or down when instructed to move right or left). We predict that such dimension errors would be common in our paradigm for the up/down and back/forward dimensions because they both involve vertical movement. If participants confuse vertical movement in one dimension with that of another dimension, then they would be especially likely to make dimension errors involving vertical movement on the computer display in the same direction (on the two-dimensional grid) so that up would be confused with forward and down would be confused with back. A third possible error type is one in which both the direction and dimension of movement are correct but the number of movements is wrong. For example, when told to move one, participants might move two or *vice versa*. This type of error would seem to be less likely when participants are explicitly told in words the number of movements to make, as in the verbal condition, than when they are forced to count the number of movements, as in the spatial condition, which would impose extra demands on working memory (Baddeley, 2007; Logie & Baddeley, 1987). Also, this type of error would seem to

be more likely when there are two required movements than when there is only one required movement because counting two movements would involve subvocalizing the number '1' as well as the number '2' if participants subvocalize a running total (see, e.g. Healy & Nairne, 1985; Logie & Baddeley, 1987).

In Experiment 2, we replicated Experiment 1 using double the number of trials. In particular, Experiment 2 included two sessions (training and testing) so that we could examine the effects of retention and transfer of the skills involved in following both types of instructions (see Healy, 2007, for a discussion of the relationship between retention and transfer). Specifically, we crossed the presentation condition used during training with the presentation condition used during testing. When the two presentation conditions are the same, performance in the second session depends only on retention of the skills used in the first session. In contrast, when the two presentation conditions switch, then performance would depend not only on retention of the skills used in the first session but also on transfer of those skills to the new skills required in the second session. Finding better performance at test when the two presentation conditions are the same than when they switch would be predicted on the basis of transfer appropriate processing (Morris, Bransford, & Franks, 1977), encoding specificity (Tulving & Thomson, 1973), procedural reinstatement (Healy, Wohldmann, & Bourne, 2005) and specificity of training (Healy, Wohldmann, Sutton, & Bourne, 2006). However, if participants eventually use the same type of dual coding for both presentation conditions, then performance when presentation conditions switch would be expected to be equivalent to performance when presentation conditions are the same. Furthermore, Experiment 2 allows us to test the hypothesis that the condition leading to worse performance at training would lead to better performance at test. This hypothesis follows from the literature showing that training under 'desirable' difficulties may enhance long-term retention and transfer (Bjork, 1994; McDaniel & Einstein, 2005; Schneider, Healy, & Bourne, 2002).

Method

Experiment 1 involved a single session, whereas Experiment 2 involved two sessions separated by a 1-week delay. In Experiment 1 and in each session of Experiment 2, half of the participants (*verbal condition*) heard navigation messages, as in previous experiments (e.g. Barshi & Healy, 2002; Schneider et al., 2004), whereas the remaining participants (*spatial condition*) heard no messages but rather saw the movements on the screen. That is, each square was highlighted in the grid of matrices in the same order that participants were to move when following the verbal instructions. The time to see the movements in the spatial condition was equated to the time to hear the instructions in the verbal condition. After all of the instructions were given, participants immediately followed the instructions by clicking in the appropriate squares with the mouse. The movements to be made by the participants were identical in the two conditions. In Experiment 2, we used all four combinations of condition in Session 1 and condition in Session 2, so that half of the participants were in the same condition in both sessions (verbal/verbal or spatial/spatial), and the other half were in different conditions in the two sessions (verbal/spatial or spatial/verbal). The verbal/verbal and spatial/spatial conditions, which included the same presentation condition each week, allowed us to examine retention, whereas the verbal/spatial and spatial/verbal conditions, which included a switch in presentation condition from training to testing, allowed us to examine transfer.

The messages in the verbal condition consisted of one to six commands telling the participants to move in the space displayed on the computer screen. The commands specified right or left movement, up or down movement and forward or back movement within the space. All of the commands consisted of three words (e.g. 'left one square') that specified the direction of movement and the number of moves. The third word was redundant with the dimension of movement; it was always the word 'square' for the right/left dimension, 'level' for the up/down dimension and 'step' for the back/forward dimension. (It was made clear to participants that verbal instructions were based on a constant orientation, so a given command always led to movements along the same direction in the grid.)

Design

A $2 \times 6 \times 6$ mixed factorial design was employed in Experiment 1. The first factor, presentation condition (verbal or spatial), was varied between participants. The second and third factors, block within session (1–6) and message length in number of commands (1–6), were varied within participants. The same design was used in Experiment 2 except that the factor of presentation condition was replaced by two crossed factors of Session 1 presentation condition (verbal or spatial) and Session 2 presentation condition (verbal or spatial).

Participants' comprehension was measured by the accuracy of their manual movements. No oral repetition (readback) responses were made (because no verbal messages were given in the spatial condition) (*cf.* Barshi & Healy, 1998, 2002). For the analysis of accuracy, each trial was scored in an all-or-none fashion (i.e. an error on any command rendered the whole trial incorrect), and we examined the proportion of correct responses.

Apparatus and materials

A new set of command sequences was created for this study in order to equate the number of movements along each movement dimension (right/left, up/down, back/forward) at each serial position of the commands. The commands themselves were used in earlier studies (e.g. Barshi & Healy, 2002). For the commands in the verbal condition, a male native English speaker had been recorded giving combinations of navigation commands in as natural a manner as possible. The speaker's voice had been digitized on a Macintosh SE30 computer using the program SoundEditPro. The clearest sample of each word (as determined both by an impressionistic judgment and by the examination of the speech waveforms) used in the commands (e.g. 'left', 'one') had then been spliced out of the natural speech stream. An iMac computer played the appropriate words according to the specific sequence indicated for each experimental trial. Thus, there was a natural stress and intonation pattern within each isolated word but not over an entire trial. The analysis of the speech waveforms also insured that each stimulus word was at least roughly comparable in amplitude or loudness. No extra pauses (i.e. periods of silence) were added between words or between commands. The average time to present the messages in the verbal condition was the same as in the spatial condition, in which the separate movements occurred at the rate of one square per second.

All participants saw four matrices stacked one above the other in the centre of the computer screen (see Figure 1 top panel). Each matrix was made of 16 squares (4×4). The starting position was given to the participants in advance of the experimental trials and was constant for the experimental trials; furthermore, it was highlighted before each message was presented in both the verbal and spatial conditions. This position was the third row

DOI: 10.1002/acp

from the top and second column from the left of the third matrix down. From any position, there were a limited number of legal moves participants could make without exiting the boundaries of the matrices. Because there were four 4×4 matrices, the only numbers possible in the commands were 1, 2 and 3. To equate the number of moves in each serial position along each dimension of movement, the numbers were limited to 1 and 2 because the three-move commands were not always possible. The commands were created such that participants would never end up outside the matrices if they followed the commands correctly. In messages of three or fewer commands, no more than one command along each movement dimension was included. To generate messages that contained more than three commands, messages had to repeat movement dimensions, with no more than two of any dimension in a message. (The repetition of a dimension is similar to what might occur when giving directions for a route along streets that include a detour. Even in flying, pilots often need to deviate from a simple flight path to avoid obstacles or weather.)

The messages varied in length from 1 to 6 commands. For example, a message with one command was ‘left one square’, and a message with six commands was ‘up one level, back one step, left one square, forward two steps, right one square, down one level’. There were an equal number of messages at each message length (1–6). The same messages were used in both conditions and in both sessions of Experiment 2. Specifically, there were 72 experimental trials divided into six 12-trial blocks. The trials were presented in a fixed pseudorandom order with the constraint that each block of 12 trials included two trials of each of the six message lengths defined in terms of the number of commands. Across the six blocks, the commands (e.g. ‘left one square’) used in a given position were employed equally often at each message length. Each of three commands along a given dimension (i.e. left one square, right one square and right two squares; down one level, up one level and up two levels; back one step, forward one step and forward two steps) was used one or two times in each position at each command message length, with each direction of movement (i.e. left, right, up, down, forward, back) used two times in each position at each command message length. Although the directions of movement were equated in this way, there were three times as many commands involving the number 1 as those involving the number 2 (and none involving the number 3). These constraints also meant that for message length 6, if there was a command to go in one direction along a dimension in the first three commands (e.g. *up*), there was always a command to go in the opposite direction in the last three commands (e.g. *down*). Note that, in addition, the constraints on legal moves meant that chance level performance for message length 1 was 1/9 (because all three possible movements along each dimension would be legal). Chance level performance would be much lower for longer message lengths; it would decrease by a factor of 1/6–1/9 (depending on which movements are legal from a given position) for each increment in message length.

Despite these constraints, which equated the different message lengths in terms of the specific commands used, there was no attempt to equate the different message lengths in terms of the ending location in the grid. However, because participants had to follow every command in a message in order to be scored as correct, they could not use the strategy of end-position seeking, in which they would consider only the final location reached. Also, finding a more efficient route to the same end location was not an appropriate objective in this task, just as it would not be an appropriate objective for a pilot given navigation instructions from an air traffic controller.

At the start of each session, participants in both conditions were given on the computer a demonstration of a trial in a given modality, including the correct responses made to a

sequence of navigation commands. The experimental trials were preceded by six practice trials in the same modality, including one of each message length. The starting point for the practice trials was a reverse image of that for the experimental trials. Specifically, it was the second row from the top and third column from the left of the second matrix down (see Figure 1 bottom panel). The order of the practice trials was chosen systematically so that message length increased across trials from 1 to 6 commands. No performance criterion was used for the practice trials, but the experimenter observed the participants during those trials and corrected them when necessary. If the participants exhibited any confusions about the fact that the commands were based on a constant orientation (i.e. they were not body centred), then the experimenter told the participants to think of the left/right movements as sliding in each direction.

Participants were shown a small three-dimensional model representing the $4 \times 4 \times 4$ computer space. The model consisted of four pieces of paper stacked one on top of another and connected at each corner with dowels. Each paper contained a 4×4 matrix similar to each 4×4 matrix shown on the computer screen. The space between the papers allowed the participants to view each matrix in its entirety. The starting point for the practice trials was filled to highlight its position. All other squares were identical and left unfilled.

Procedure

Participants heard (verbal) or saw (spatial) a sequence of movements of different message lengths (i.e. different numbers of commands). In both conditions, the sequence of movements was followed by a beep. After the beep, the starting location was highlighted (again). All participants followed the sequence of movements they had heard or seen by clicking with the computer mouse in the appropriate places on the grid. Participants had to click every square they passed, so, for example, if the first spoken command was 'right 2', they had to click on the square immediately to the right of the starting position and then on the square immediately to the right of that. The number of clicks had to equal the number spoken in the verbal condition (i.e. 1 = one click, 2 = two clicks) or the number of moves shown in the spatial condition. After following the sequence of movements, the participants had to click the DONE button (see Figure 1 bottom panel). There was a 2-second pause between trials.

The instructions on the procedure for the verbal and spatial conditions were identical except with respect to the modality in which the sequence of movements was presented. With these instructions, all participants were shown the three-dimensional model of the space and were told that the computer display represented that model, as in the previous experiments (e.g. Barshi & Healy, 2002).

At the conclusion of the experiment, participants received a questionnaire asking them how they coded the information both while receiving the stimuli and while making the movement responses. In particular, for the spatial condition participants were asked, 'As you watched the sequences did you translate the movement sequences into words?' On the other hand, for the verbal condition participants were asked, 'As you listened to the instructions did you translate the words into movement sequences?' In addition, for both conditions participants were asked, 'As you followed the sequence of movements, clicking in each of the spaces, did you hear words in your mind, see words in your mind, or visualize the path in your mind?' (For the verbal condition the word 'instructions' replaced the phrase 'sequence of movements'.)

Although the post-experiment questionnaire was administered only at the end of Session 2 in Experiment 2, participants were asked about the strategies they used in Session 1 as

DOI: 10.1002/acp

well as in Session 2. The questionnaire was not administered at the end of Session 1 to avoid any bias on Session 2 performance that might have been caused by having previously answered the questions.

Participants

Thirty-two undergraduate students at the University of Colorado, Boulder, participated in Experiment 1, and 48 participated in Experiment 2 for credit in a course in introductory psychology. There were 16 participants in each of the two conditions (verbal, spatial) of Experiment 1 and 12 participants in each of the four conditions (verbal/verbal, spatial/spatial, verbal/spatial, spatial/verbal) of Experiment 2. Participants were assigned to conditions according to a fixed rotation on the basis of time of arrival for testing. All participants were native speakers of English, and no participant was in both experiments.

Results and discussion

Training

We report first the results of training, which involves the single session in Experiment 1 and Session 1 in Experiment 2.

As in previous studies (e.g. Barshi & Healy, 2002; Schneider et al., 2004), the proportion of correct responses declined dramatically as message length increased; the main effect of message length was significant, Experiment 1: $F(5, 150) = 243.73$, MSE $= 0.088$, $p < .01$; Experiment 2: $F(5, 230) = 324.95$, MSE $= 0.086$, $p < .01$. Performance declined rapidly across message lengths, with the largest drops between message lengths 3 and 4 and 4 and 5 (see Figure 2 top and middle panels). The decline was the same for the verbal and spatial conditions; the interaction of message length and condition was not significant. Also significant was the main effect of block, Experiment 1: $F(5, 150) = 5.07$, MSE $= 0.059$, $p < .01$; Experiment 2: $F(5, 230) = 6.67$, MSE $= 0.068$, $p < .01$, reflecting an improvement in proportion of correct responses only across the first three or four blocks of trials (Experiment 1: Block 1, $M = .570$, SEM $= .030$; Block 2, $M = .594$, SEM $= .030$; Block 3, $M = .617$, SEM $= .031$; Block 4, $M = .677$, SEM $= .028$; Block 5, $M = .651$, SEM $= .030$; Block 6, $M = .643$, SEM $= .030$; Experiment 2: Block 1, $M = .575$, SEM $= .025$; Block 2, $M = .608$, SEM $= .025$; Block 3, $M = .670$, SEM $= .024$; Block 4, $M = .656$, SEM $= .024$; Block 5, $M = .658$, SEM $= .023$; Block 6, $M = .672$, SEM $= .023$).

Although there was no overall difference between the spatial and verbal conditions, during the first three blocks of the session participants were more accurate in the spatial condition than in the verbal condition, but the two conditions were equivalent in the last three blocks (see Figure 3); the interaction of presentation condition and block was significant, Experiment 1: $F(5, 150) = 3.16$, MSE $= 0.059$, $p < .01$; Experiment 2: $F(5, 230) = 3.03$, MSE $= 0.068$, $p = .01$. The initial advantage for the spatial condition over the verbal condition was an unexpected finding, especially because having words might have helped participants use dual coding based on a representation that is verbal as well as spatial (see, e.g. Paivio, 1991, 2007). However, there are at least four reasons why performance might be worse in the verbal condition than in the spatial condition. First, by the S-C-R compatibility model, spatial tasks are more compatible with visual than with auditory inputs (Wickens et al., 1984). Second, experience is required for participants to form spatial mental models given verbal route instructions (e.g. Brunyé & Taylor, 2008a, b). Third, a three-dimensional spatial representation might be more likely to be formed given verbal than given spatial instructions. Because the response required of participants

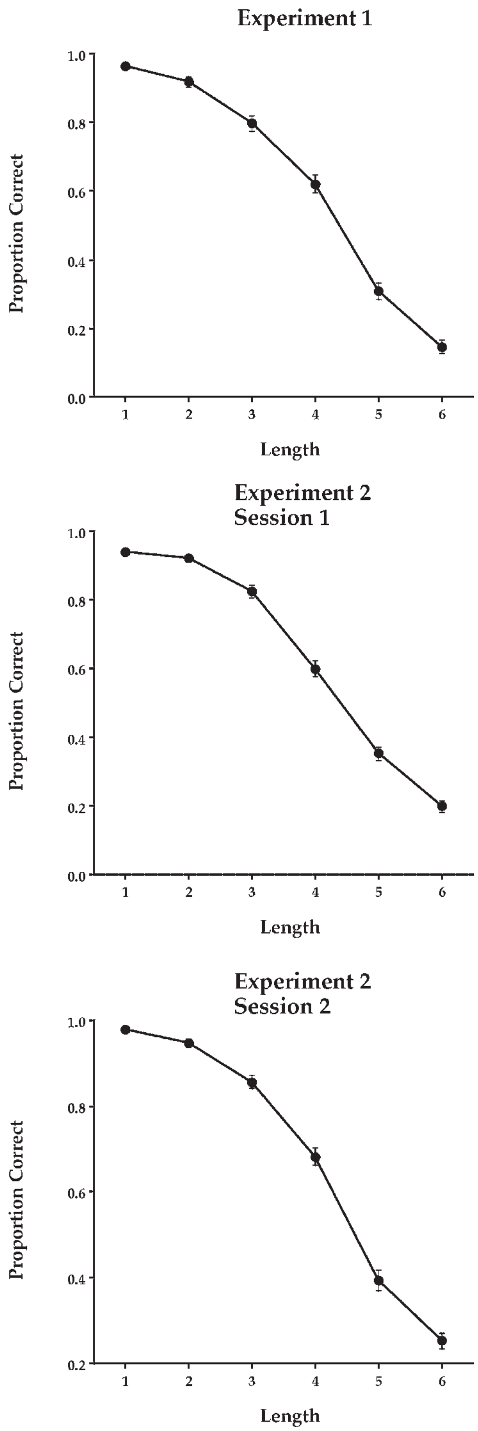


Figure 2. Proportion correct as a function of message length in Experiment 1 (top panel), Session 1 of Experiment 2 (middle panel) and Session 2 of Experiment 2 (bottom panel). Bars show standard errors of the mean

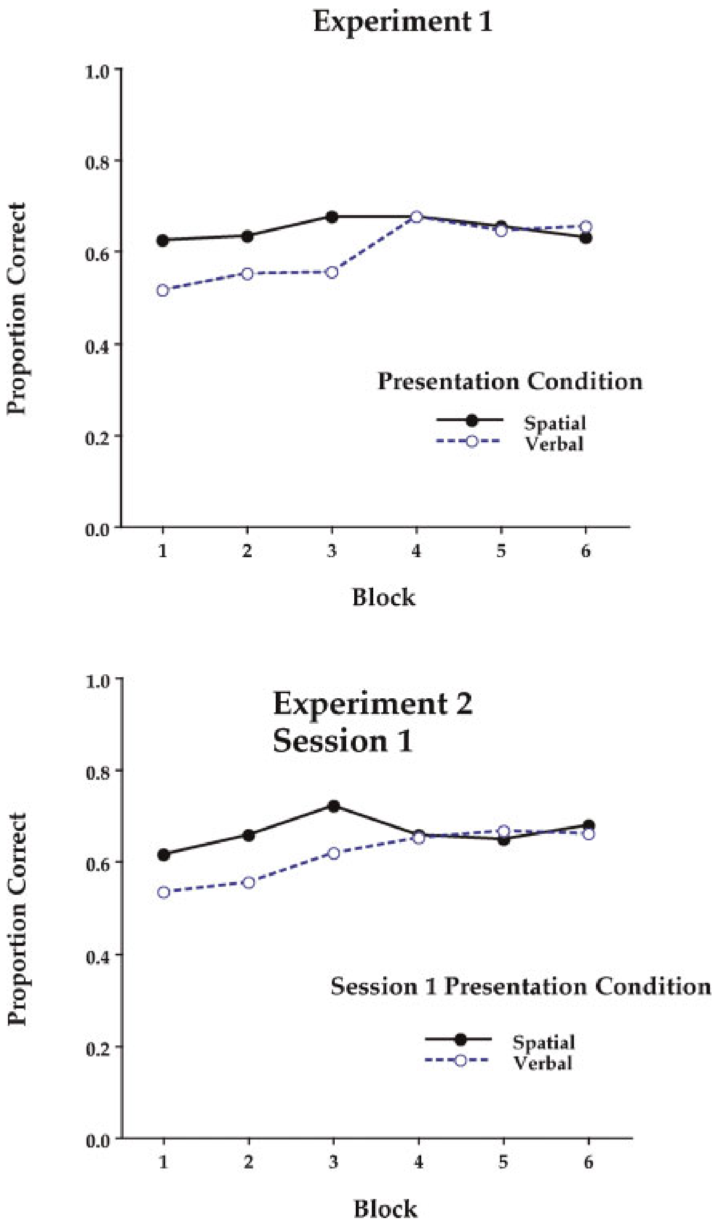


Figure 3. Proportion correct as a function of presentation condition and block in Experiment 1 (top panel) and in Session 1 of Experiment 2 (bottom panel)

involves navigating in a two-dimensional space, any three-dimensional representation would require an extra translation process. Fourth, participants in the verbal condition, but not those in the spatial condition, might undergo confusions due to the fact that the verbal instructions were based on a constant orientation, and it might require some training to overcome that confusion if the verbal commands were interpreted as body centred.

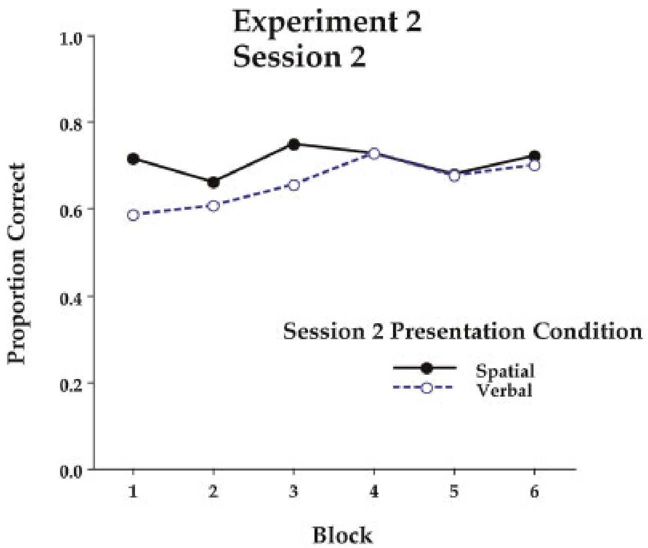
Retention and transfer

Session 2 of Experiment 2. For Session 2 of Experiment 2, we included the variable of Session 2 presentation condition as well as the variable of Session 1 presentation condition. Again, there was no overall difference between the spatial and verbal conditions, both in terms of the Session 1 condition and the Session 2 condition. The results for Session 2 of Experiment 2 were very similar to those for Session 1. As for Session 1, in the first three blocks of Session 2, participants showed a higher proportion of correct responses in the spatial condition than in the verbal condition, but the two conditions were equivalent in the last three blocks (see Figure 4 top panel); the interaction of Session 2 condition and block was significant, $F(5, 220) = 3.08$, $\mathrm{MSE} = 0.064$, $p = .01$. Also, as previously, the proportion of correct responses decreased sharply as message length increased; the main effect of message length was significant, $F(5, 220) = 214.77$, $\mathrm{MSE} = 0.122$, $p < .01$ (see Figure 2 bottom panel). Furthermore, the main effect of block was significant, $F(5, 220) = 6.02$, $\mathrm{MSE} = 0.064$, $p < .01$, although there is not a consistent pattern of improvement (Block 1, $M = .651$, $\mathrm{SEM} = .024$; Block 2, $M = .635$, $\mathrm{SEM} = .025$; Block 3, $M = .703$, $\mathrm{SEM} = .024$; Block 4, $M = .729$, $\mathrm{SEM} = .022$; Block 5, $M = .679$, $\mathrm{SEM} = .024$; Block 6, $M = .712$, $\mathrm{SEM} = .022$).

During Session 2 of Experiment 2, Session 1 condition showed the reverse pattern across blocks as in Session 1. Specifically, in Session 2, participants' initial performance was more accurate when they had received verbal presentation in Session 1 rather than spatial presentation, but after three blocks, the two Session 1 presentation conditions were equivalent (see Figure 4 bottom panel); the interaction of Session 1 condition and block was significant, $F(5, 220) = 2.84$, $\mathrm{MSE} = 0.064$, $p = .02$. The disadvantage for training in the spatial condition can be attributed to the fact that such training does not provide any learning of how to translate the verbal message into movements. These results are consistent with the hypothesis raised earlier that providing more difficult training (verbal, which requires translation from a verbal to a spatial representation) would lead to superior performance at test relative to providing less difficult training (spatial, which requires no translation). This hypothesis derives from earlier studies showing that training under 'desirable' difficulties, although slowing the rate of acquisition, enhances long-term retention and transfer (Bjork, 1994; McDaniel & Einstein, 2005; Schneider et al., 2002).

There was also a significant interaction of Session 1 and Session 2 presentation conditions, $F(1, 44) = 4.87$, $\mathrm{MSE} = 0.575$, $p = .03$, reflecting the fact that the proportion of correct responses was higher when there was a match in Session 1 and Session 2 conditions (i.e. when only retention was required; spatial/spatial, $M = .742$, $\mathrm{SEM} = .018$; verbal/verbal, $M = .708$, $\mathrm{SEM} = .019$) than when the condition in Session 1 did not match the condition in Session 2 (i.e. when transfer was also required; spatial/verbal, $M = .611$, $\mathrm{SEM} = .020$; verbal/spatial, $M = .678$, $\mathrm{SEM} = .020$). This advantage for matching Session 1 condition and Session 2 condition was more prominent at the longer message lengths than at the shorter message lengths (see Figure 5); there was a significant three-way interaction of Session 1 condition, Session 2 condition, and message length, $F(5, 220) = 2.69$, $\mathrm{MSE} = 0.122$, $p = .02$. These findings are in agreement with the literature demonstrating transfer appropriate processing (Morris et al., 1977), encoding specificity (Tulving & Thomson, 1973), procedural reinstatement (Healy et al., 2005) and specificity of training (Healy et al., 2006).

Comparison of Sessions 1 and 2 of Experiment 2. We conducted two additional analyses to focus on retention and transfer effects, respectively, in Experiment 2. The retention



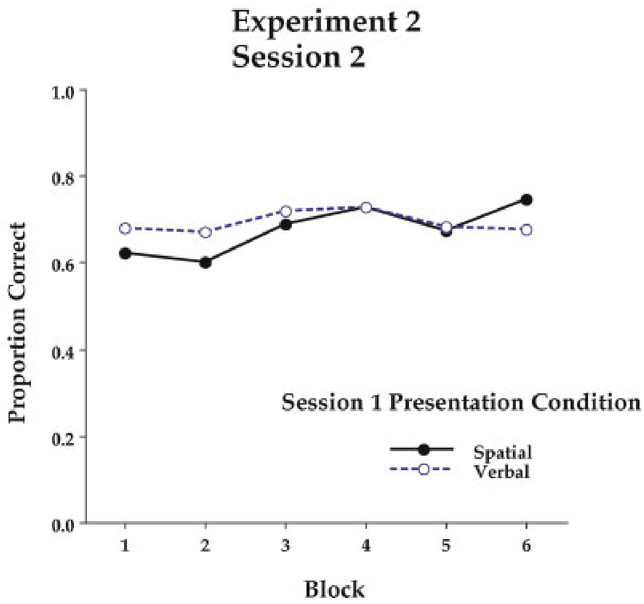


Figure 4. Proportion correct as a function of Session 2 presentation condition and block in Session 2 of Experiment 2 (top panel). Proportion correct as a function of Session 1 presentation condition and block in Session 2 of Experiment 2 (bottom panel)

analysis was restricted to those groups of participants who received the same condition in both sessions; it included the factors of session, condition and block. This analysis allowed us to examine whether improvements in performance achieved during the first session were maintained across the 1-week delay and whether a superiority for verbal instructions emerged with more practice. We found a significant main effect of session, $F(1, 22) = 19.96$, $MSE = 0.055$, $p < .01$, reflecting the fact that the proportion of correct

DOI: 10.1002/acp

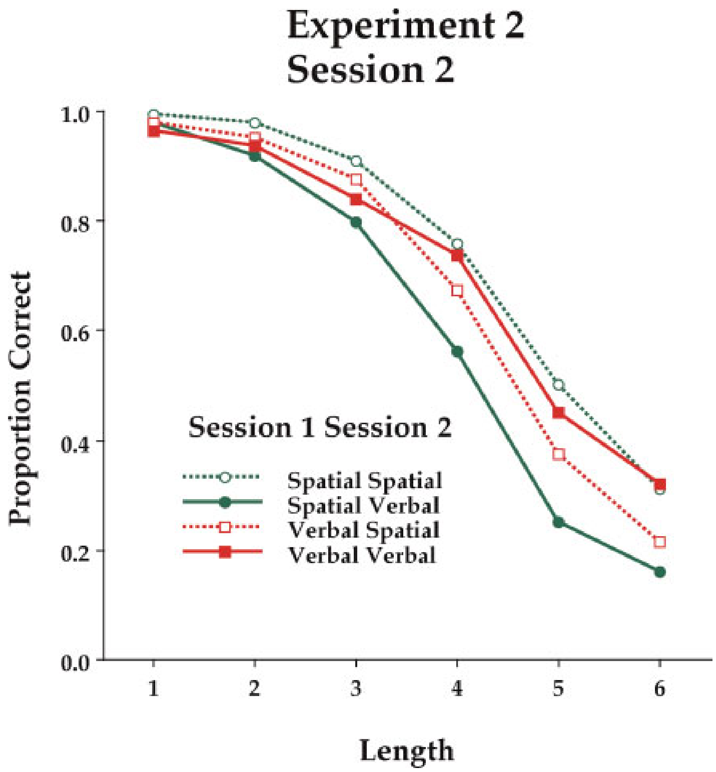


Figure 5. Proportion correct as a function of message length, Session 1 condition and Session 2 condition in Session 2 of Experiment 2

responses was higher overall in Session 2 ($M = .725$, SEM = .013) than in Session 1 ($M = .675$, SEM = .013), demonstrating the benefits of training. There was also a main effect of block, $F(5, 110) = 3.63$, MSE = 0.085, $p < .01$, because of general improvements across blocks (see Figure 6 top panel). Furthermore, the interaction of session and condition was significant, $F(1, 22) = 4.43$, MSE = 0.055, $p < .05$, because the advantage for the spatial condition was larger in Session 1 than in Session 2. Nevertheless, it is clear from Figure 6 (top panel) that no superiority for the verbal condition emerges, even at the end of the second session. It is also interesting to note that for both conditions there was no loss in the proportion of correct responses across the 1-week delay spanning the two sessions; thus, there was perfect retention of the acquired ability to perform this task successfully.

The transfer analysis was restricted to those groups of participants who received different conditions in the two sessions; it included the factors of session, Session 1/Session 2 presentation condition and block. This analysis allowed us to examine whether the acquired ability to perform the task under one condition can be transferred to that under the other condition. As in the previous analysis on participants who were in the same condition in both sessions, the present analysis on participants who were in different conditions across sessions showed a general improvement across the initial blocks in each session; the main effect of block was significant, $F(5, 110) = 8.50$, MSE = 0.063, $p < .01$ (see Figure 6 bottom panel). More importantly, there was a three-way interaction of block, session and

DOI: 10.1002/acp

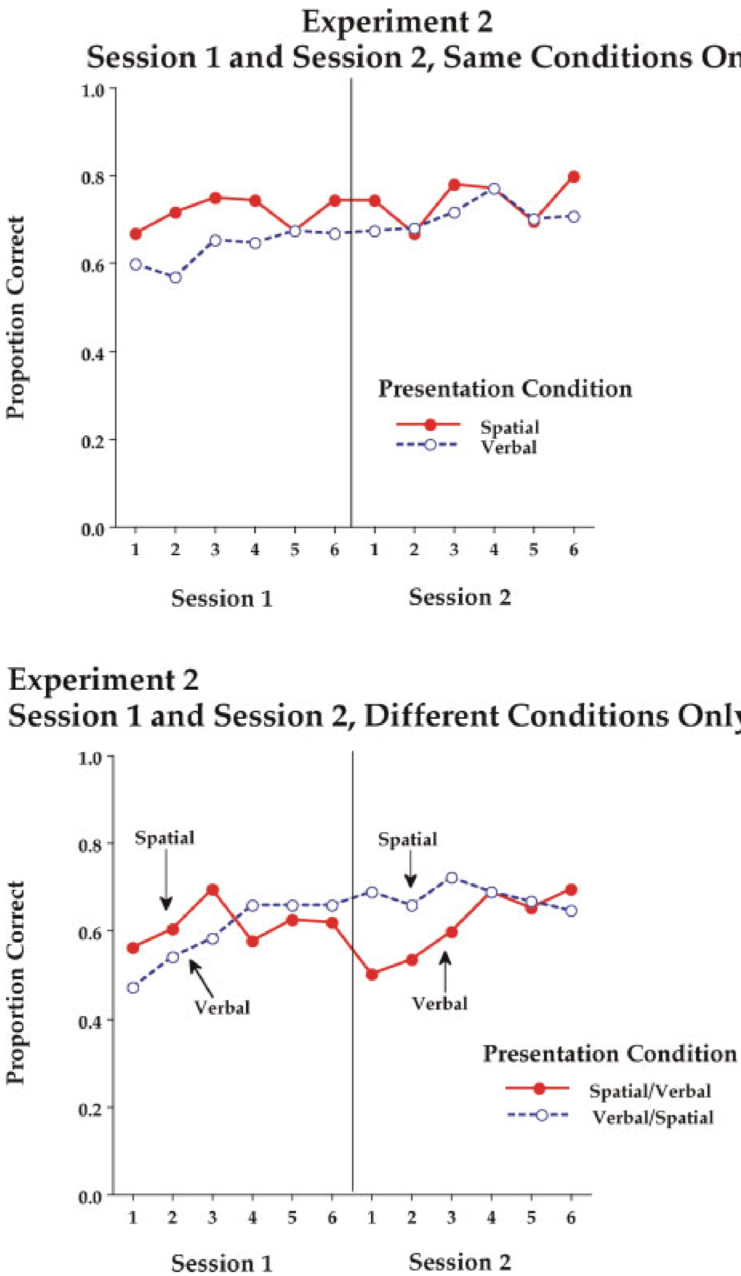


Figure 6. Proportion correct for *same* conditions only (top panel) and *different* conditions only (bottom panel) as a function of Session 1/Session 2 presentation condition and block in both sessions of Experiment 2

Session 1/Session 2 presentation condition, $F(5, 110) = 7.73$, $MSE = 0.066$, $p < .01$, reflecting the fact that there was essentially perfect transfer from the verbal to the spatial condition (with performance at the start of Session 2 in that case higher than that for participants in the spatial condition at the start of Session 1) but essentially no transfer from

DOI: 10.1002/acp

the spatial to the verbal condition (with performance at the start of Session 2 more similar in that case to that for participants in the verbal condition at the start of Session 1). The finding that there was virtually perfect transfer from the verbal to the spatial condition but virtually no transfer from the spatial to the verbal condition provides further evidence for the hypothesis that difficult training is desirable in enhancing subsequent transfer.

Figure 6 (bottom panel) also reveals that in both sessions, participants showed a higher proportion of correct responses in the spatial condition than in the verbal condition during the first three blocks of practice but essentially no difference between the two conditions in the last three practice blocks. These findings are consistent with the earlier observation that whenever participants begin a new presentation condition it takes about three blocks for the verbal condition to catch up to the spatial condition. The findings are also consistent with the hypothesis that participants in the verbal condition learn across the first three blocks how to translate the verbal instructions into movement sequences.

Retrospective reports

Most participants in the spatial condition responded in the retrospective report questionnaire that they translated the movement sequences into words (11 of 16 participants in Experiment 1, 15 of 24 participants in Session 1 of Experiment 2 and 20 of 24 participants in Session 2 of Experiment 2) or heard words in their minds (8 of 16 participants in Experiment 1, 13 of 24 participants in Session 1 of Experiment 2 and 16 of 24 participants in Session 2 of Experiment 2).

Most participants in the verbal condition reported that they translated the words into movements as they heard the instructions (15 of 16 participants in Experiment 1, all 24 participants in Session 1 of Experiment 2 and 23 of 24 participants in Session 2 of Experiment 2) or that they visualized the path as they followed the instructions (13 of 16 participants in Experiment 1, 17 of 24 participants in Session 1 of Experiment 2 and 21 of 24 participants in Session 2 of Experiment 2). Thus, most participants in the two presentation conditions said that they used both verbal and spatial representations, thereby implying the use of dual coding for both presentation formats (e.g. Paivio, 1991, 2007) and suggesting that participants may be able to transfer from one presentation format to the other. This suggestion is apparently inconsistent with the strong advantage for the spatial presentation format during the initial blocks of each session. However, this suggestion might not actually be inconsistent with this finding because the retrospective reports were collected only at the end of the experiment, but the differences between presentation formats occurred only in the first half of each session. Presumably during the initial blocks, participants in the verbal condition learned how to translate the verbal instructions into a sequence of movements equivalent to that shown to participants in the spatial condition. Such a translation of verbal commands into spatial movements is consistent with expectations derived from earlier studies (e.g. Barshi & Healy, 2002) suggesting that the verbal representation of navigation instructions is influenced by the spatial representation. The underlying nature of the verbal and spatial representations is illuminated to some extent by the different patterns of initial errors made by participants in the two presentation conditions.

Initial errors

Movement dimension. To elucidate possible differences between verbal and spatial conditions in the movement dimensions along which errors are made, we examined the first error made on a given trial and counted the number of errors (of all types) made along each

movement dimension. We analysed these errors separately for each session in Experiment 2. In the analysis of these errors for Experiment 1 and for Session 1 of Experiment 2, we included only the factors of (Session 1) presentation condition (verbal, spatial) and correct movement dimension (right/left, up/down, back/forward). For the analysis of Session 2 of Experiment 2, we also included the factor of Session 2 condition. The main effect of movement dimension was significant, Experiment 1: $F(2, 60) = 9.93$, MSE = 12.824, $p < .01$; Experiment 2 Session 1: $F(2, 92) = 15.38$, MSE = 15.950, $p < .01$; Experiment 2 Session 2: $F(2, 88) = 7.10$, MSE = 10.687, $p < .01$. The largest mean number of initial errors was on the up/down and back/forward dimensions, and the smallest was on the right/left dimension (see Table 1).

The interaction of condition and movement dimension was also significant in Experiment 1 and in Session 2 of Experiment 2 (for Session 2 condition), although not in Session 1 of Experiment 2 [Experiment 1: $F(2, 60) = 3.27$, MSE = 12.824, $p = .04$; Experiment 2 Session 1: $F(2,92) = 1.25$, MSE = 15.950, $p = .29$; Experiment 2 Session 2: $F(2, 88) = 4.25$, MSE = 10.687, $p = .02$], because the spatial condition showed a larger effect of movement dimension than did the verbal condition (see Figure 7). More specifically, the biggest difference in each case (even in Session 1 of Experiment 2) between the spatial and verbal conditions in initial errors was along the right/left dimension, with fewer errors in the spatial than in the verbal condition. This reduction of errors along the right/left dimension in the spatial condition relative to the verbal condition could result from one of four factors: (a) The difference between right and left could be more obvious in the visual display than the difference between up and down or that between back and forward. However, that factor is unlikely because the up/down dimension involves the greatest distance, extending from one matrix to another. (b) The right/left dimension could be the most distinctive in the visual display. This factor seems more likely because the up/down and back/forward dimensions both involve vertical movement in the two-dimensional screen representation, whereas only the right/left dimension involves horizontal movement. (c) The participants may not know how to give distinctive and consistent verbal labels to the up/down and back/forward movements in the spatial condition. (d) Participants in the verbal condition might have some difficulty mapping the words ‘right’ and ‘left’ to appropriate moves and such difficulty is eliminated when no words are spoken (see, e.g. Corballis & Beale, 1970). Difficulty in using the words ‘right’ and ‘left’ might result from the fact that the words ‘right’ and ‘left’ may be

Table 1. Mean (and standard error of the mean) number of initial errors as a function of dimension

Experiment and session	Errors of all types	Dimension errors
Experiment 1		
Up/down	10.562 (0.685)	5.875 (0.587)
Back/forward	8.219 (0.704)	4.531 (0.544)
Right/left	6.594 (0.638)	3.750 (0.381)
Experiment 2 Session 1		
Up/down	9.708 (0.678)	5.250 (0.463)
Back/forward	9.396 (0.714)	4.583 (0.493)
Right/left	5.646 (0.495)	2.917 (0.297)
Experiment 2 Session 2		
Up/down	7.979 (0.548)	5.062 (0.472)
Back/forward	8.042 (0.637)	4.021 (0.381)
Right/left	5.833 (0.629)	3.042 (0.407)

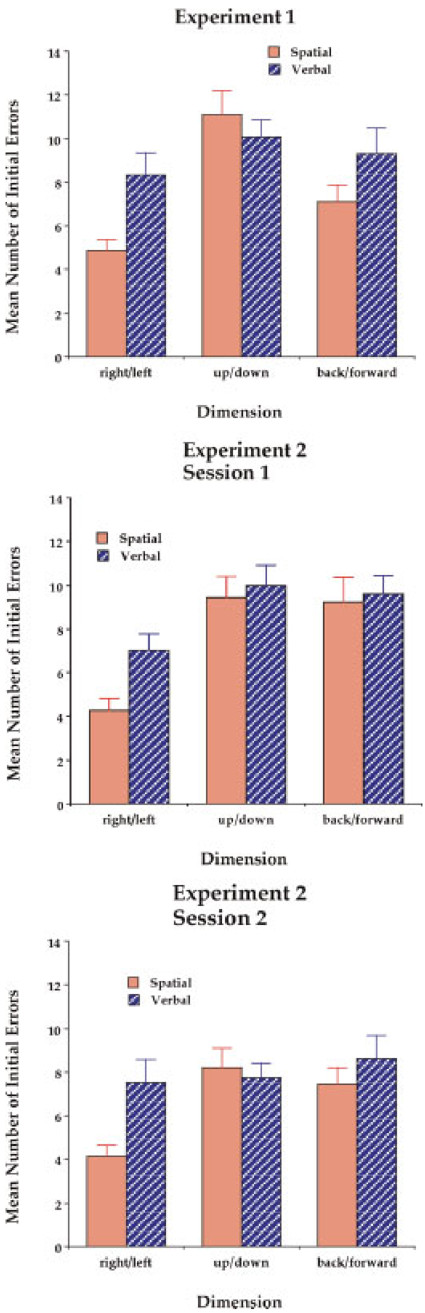


Figure 7. Mean number of initial errors (of all types) as a function of movement dimension in Experiment 1 for the verbal and spatial conditions (top panel), in Session 1 of Experiment 2 for the verbal and spatial Session 1 conditions (middle panel), and in Session 2 of Experiment 2 for the verbal and spatial Session 2 conditions (bottom panel). Bars show standard errors of the mean

DOI: 10.1002/acp

confusable because they differ for the speaker and the listener whenever the two individuals are facing each other. Furthermore, as Farrell (1979) remarked, understanding the words ‘left’ and ‘right’ requires the ability to ‘tell’ left from right, whereas copying movements does not, and the ability to tell left from right is not trivial (Corballis & Beale, 1970). It is also worth noting that the finding of a decrease in right/left errors in a non-verbal condition relative to a verbal condition is consistent with a study by Maki (1979) involving spatial judgments (but inconsistent with the findings of Farrell, 1979), although right/left judgments were worse than up/down judgments in the earlier verbal conditions of these studies. Perhaps up/down movements were worse than right/left movements in our study but not in the studies by Farrell and Maki because only in our study were there two different forms of vertical movement (up/down and back/forward), thereby making up/down movement more prone to errors.

If participants in the verbal condition change from one type of coding in the first half of the session (i.e. in the first three blocks) to a new type of coding in the second half of the session (i.e. in the last three blocks) and if the new type of coding is like that used in the spatial condition, then we might expect to see a reduction in the number of errors along the right/left dimension in the second half of the session relative to the first half of the session for those participants. A *post hoc* examination of initial errors (of all types) as a function of movement dimension was not in agreement with this prediction. Specifically, in Experiment 1, the mean number of errors along the right/left dimension in the verbal condition did not decrease from the first half ($M = 4.062$, SEM $= 0.574$) to the second half ($M = 4.250$, SEM $= 0.642$) of the session, despite the fact that there was such a decrease in the mean number of errors both for errors along the up/down dimension (first half: $M = 5.938$, SEM $= 0.574$; second half: $M = 4.125$, SEM $= 0.446$) and for errors along the back/forward dimension (first half: $M = 5.938$, SEM $= 0.727$; second half: $M = 3.375$, SEM $= 0.625$). A similar pattern was evident for the first session of Experiment 2. In this case, for the verbal condition there was no change at all in the mean number of errors along the right/left dimension from the first half ($M = 3.500$, SEM $= 0.426$) to the second half ($M = 3.500$, SEM $= 0.485$) of the session. Again, however, there was a decline in the mean number of errors along the up/down dimension (first half: $M = 5.792$, SEM $= 0.511$; second half: $M = 4.208$, SEM $= 0.485$) and along the back/forward dimension (first half: $M = 5.583$, SEM $= 0.645$; second half: $M = 4.000$, SEM $= 0.413$). Although this finding is inconsistent with the predicted pattern of errors, an examination of a more specific type of initial error, a direction error, is consistent with the prediction.

Direction errors. A count was made of the number of times the first error consisted of a direction error, in which participants responded with a movement along the correct dimension and made the correct number of movements but moved in the wrong direction (e.g. they were supposed to move left one but moved right one instead). In Experiment 1 and in both sessions of Experiment 2, all of the errors of this type along the right/left dimension were made in the verbal condition; the mean number of errors of this type was thus 0 for the spatial condition in each case and was 0.812 (SEM $= 0.440$), 0.542 (SEM $= 0.190$) and 0.708 (SEM $= 0.221$) for the verbal condition in Experiment 1, Experiment 2 Session 1 and Experiment 2 Session 2, respectively. This finding supports the hypothesis that participants in the verbal condition find it difficult to map the words ‘left’ and ‘right’ to the appropriate movements (Corballis & Beale, 1970; Farrell, 1979).

A *post hoc* examination of these initial right/left direction confusions in the verbal condition reveals a decline in errors from the first half of the session to the second half of

the session in Experiment 1 (first half $M = 0.688$, SEM $= 0.373$; second half $M = 0.125$, SEM $= 0.085$) as well as in both the first session of Experiment 2 (first half $M = 0.375$, SEM $= 0.132$; second half $M = 0.167$, SEM $= 0.098$) and the second session of Experiment 2 (first half $M = 0.417$, SEM $= 0.146$; second half $M = 0.292$; SEM $= 0.112$).

Dimension errors. A second count was made of the number of times the first error consisted of a dimension error, in which participants responded with the correct number of movements but moved along an incorrect dimension (e.g. they were supposed to go up one but they went forward one instead). For Experiment 1 and both sessions of Experiment 2, the mean number of errors of this type was higher when the correct dimension was up/down or back/forward than when it was right/left (see Table 1). As predicted, the number of dimension errors was especially high when both dimensions involved vertical movement and the movement was in the same direction. Specifically, the mean number of errors was very high when up was the prescribed dimension and the participant went forward instead (Experiment 1: $M = 2.219$, SEM $= 0.329$; Experiment 2 Session 1: $M = 2.354$, SEM $= 0.280$; Experiment 2 Session 2: $M = 2.083$, SEM $= 0.273$) and when down was the prescribed dimension and the participant went back instead (Experiment 1: $M = 1.969$, SEM $= 0.248$; Experiment 2 Session 1: $M = 1.708$, SEM $= 0.217$; Experiment 2 Session 2: $M = 1.729$, SEM $= 0.178$); no other dimension errors were as frequent as these, and these two types of errors were the most frequent in both the verbal condition (Experiment 1: up/forward: $M = 1.375$, SEM $= 0.202$; down/back: $M = 1.625$, SEM $= 0.375$; Experiment 2 Session 1: up/forward: $M = 2.083$, SEM $= 0.371$; down/back: $M = 1.458$, SEM $= 0.289$; Experiment 2 Session 2: up/forward: $M = 1.625$, SEM $= 0.254$; down/back: $M = 1.750$, SEM $= 0.264$) and the spatial condition (Experiment 1: up/forward: $M = 3.062$, SEM $= 0.559$; down/back: $M = 2.312$, SEM $= 0.312$; Experiment 2 Session 1: up/forward: $M = 2.625$, SEM $= 0.421$; down/back: $M = 1.958$, SEM $= 0.321$; Experiment 2 Session 2: up/forward: $M = 2.542$, SEM $= 0.470$; down/back: $M = 1.708$, SEM $= 0.244$). Both of these cases reflect confusion between the up/down and back/forward dimensions when the participant was supposed to go up or down from one matrix to another, but instead went forward or back within a single matrix.

Movement magnitude. Another potential difference between the verbal and spatial conditions involves errors in which the movement magnitude is too short or too long. Specifically, participants given a command requiring them to make two clicks might make only one, or those given a command to make one click might make two instead. Because there were three times as many commands involving one required click as those involving two required clicks, participants might be biased to respond with only a single click whenever they are uncertain of the number. Again, we examined the first error made on a given trial, and this time counted the number of errors made along the correct dimension and in the correct direction but with the wrong magnitude. In the analysis of these errors in Experiment 1 and in Session 1 of Experiment 2, we included only the factors of presentation condition (verbal, spatial) and number of required clicks (1, 2). In the comparable analysis of Session 2 of Experiment 2, we included the factors of Session 1 condition (verbal, spatial) and Session 2 condition (verbal, spatial) as well as number of required clicks. For Experiment 1, the only significant effect was the interaction of the two factors, Experiment 1: $F(1, 30) = 7.72$, MSE $= 5.062$, $p = .01$; which was also marginally significant in Session 1 of Experiment 2, $F(1, 46) = 3.32$, MSE $= 4.291$, $p = .07$, as was the interaction of Session 2 condition and number of required clicks in Session 2 of Experiment 2, $F(1, 44) = 3.31$, MSE $= 5.289$, $p = .08$. For Experiment 2 Sessions 1 and 2, but not for

Experiment 1, there was also a significant main effect of the number of required clicks [Experiment 1, $F(1, 30) = 3.57$, MSE = 5.062, $p = .07$; Experiment 2 Session 1: $F(1, 46) = 7.89$, MSE = 4.291, $p < .01$; Experiment 2 Session 2: $F(1, 44) = 4.35$, MSE = 5.289, $p = .04$]. These effects reflect a larger mean number of errors when the correct response required two clicks (Experiment 1: $M = 4.500$, SEM = 0.492; Experiment 2 Session 1: $M = 4.792$, SEM = 0.521; Experiment 2 Session 2: $M = 4.292$, SEM = 0.533) than when it required only one click (Experiment 1: $M = 1.875$, SEM = 0.328; Experiment 2 Session 1: $M = 2.833$, SEM = 0.469; Experiment 2 Session 2: $M = 2.458$, SEM = 0.421) in the spatial condition (despite the fact that there were three times as many directions involving one click as involving two clicks and, thus, three times as many opportunities to make an error involving one click as involving two clicks) but a smaller difference (not always in the same direction) in the verbal condition (Experiment 1: two clicks $M = 3.625$, SEM = 0.539; one click $M = 4.125$, SEM = 0.821; Experiment 2 Session 1: two clicks $M = 3.875$, SEM = 0.475; one click $M = 3.458$, SEM = 0.518; Experiment 2 Session 2: two clicks $M = 3.667$, SEM = 0.473; one click $M = 3.542$, SEM = 0.565).

Thus, in Experiment 1 and both sessions of Experiment 2, participants in the spatial condition (who saw movements on the screen) made more initial movement magnitude errors when two clicks were required (so two movements were shown) than when one click was required (so only a single movement was shown), with little difference between two and one required clicks for participants in the verbal condition, although the number of movement magnitude errors is roughly equivalent overall in the two conditions. With respect to the comparison of verbal and spatial conditions, there were more initial movement magnitude errors in the spatial condition than in the verbal condition when the correct response required two clicks, but the opposite was true when the correct response required one click. In the spatial condition the movement magnitude errors when two clicks are required might reflect the fact that participants are not given explicit information about the number of movements, as they are in the verbal condition (e.g. they are told 'forward two'). Participants in the spatial condition must derive the number of required movements by counting them, whereas no counting is required in the verbal condition, and counting requires working memory resources (Baddeley, 2007; Logie & Baddeley, 1987). This explanation relies on the assumption that counting should lead to more confusion with two movements than with one movement, presumably because counting two movements involves saying 'one' as well as 'two' (see, e.g. Healy & Nairne, 1985).

The lack of explicit information about the number of required clicks in the spatial condition (which imposes an extra working memory load of counting), though, cannot explain why performance in the spatial condition is actually better than that in the verbal condition in the first three blocks, why the total number of initial errors due to discrepancies in movement magnitude does not significantly differ between the two conditions, or why the number of initial movement magnitude errors with one required click was higher in the verbal condition than in the spatial condition.

Conclusion

We compared performance on following navigation instructions presented verbally (as auditory commands) or spatially (as a visual sequence of movements). At the outset, we considered three possible outcomes of this comparison. First, if participants use the same strategies in the verbal and spatial conditions, then spatial presentation might yield comparable performance to that found with verbal presentation so instructions in the two modalities would show the same pattern of results. Second, on the basis of the assumption

DOI: 10.1002/acp

that effort would be required to create a spatial representation from a verbal representation, spatial presentation might yield better performance than verbal presentation (see also Wickens et al., 1984, for a similar prediction based on a model of S-C-R compatibility). Third, on the basis of a dual-coding hypothesis (e.g. Paivio, 1991, 2007), according to which two memory codes would be more likely to be formed with verbal than with spatial presentation, verbal presentation might yield better performance than would spatial presentation. We found support for the second of these possible outcomes because performance was better with the spatial presentation condition than with the verbal presentation condition, but only for the first half of each session. However, there was also some support for the third outcome because in Experiment 2 there was an advantage during the first half of Session 2 for having trained in the verbal condition in Session 1.

In each session of the two present experiments, it appears that participants in the verbal condition learned across blocks how to translate the verbal message into movements, so that after the first half of each session, performance in the verbal condition was equivalent to that in the spatial condition, in which no words were spoken. That is, presenting the instructions in a spatial display led to better performance in the first half of the session than did presenting the instructions as auditory commands. It is interesting to note that the three-block advantage for the spatial condition was found both at the start of training in the two experiments and at the start of transfer for those participants in Experiment 2 who were exposed to different conditions in the two sessions.

For participants who were in the same condition in both sessions of Experiment 2, a superiority for verbal instructions did not emerge as trials progressed; instead the two conditions were roughly equivalent in the second session. This finding suggests that the only change in coding strategy occurs during the first three blocks of a given session, with no subsequent changes. Furthermore, performance when conditions were switched across sessions was not equivalent to performance when conditions were the same in the two sessions in Experiment 2, suggesting that participants do not use the same type of dual coding for both presentation conditions at least in the first half of the session, despite their retrospective reports indicating that they use both verbal and spatial representations for both presentation conditions. It is possible, however, that participants in the verbal condition changed from a single verbal code in the first half of the session to dual verbal and spatial coding in the second half of the session, and the retrospective reports might have reflected only the coding used in the second half of the session.

The differences between the coding used for the verbal and spatial presentation conditions are illuminated by analyses of the first errors (of all types) made on a given trial as a function of movement dimension. In both experiments, the advantage for the spatial condition relative to the verbal condition seems to be due in part to the fact that there were fewer errors along the right/left dimension for the spatial condition than for the verbal condition. However, the *post hoc* analyses of each experiment showed that the mean number of errors along the right/left dimension in the verbal condition did not decrease from the first half to the second half of the session, even though there was such a decrease for errors along the up/down and back/forward dimensions.

The *post hoc* examination supports the conclusion that the form of coding used in the verbal condition does improve from the first to the second half of the session but not because it changes to that used in the spatial condition. Thus, participants' representation of movements in the verbal condition is not identical to participants' representation of movements in the spatial condition, even with sufficient practice that allows participants in the verbal condition to perform at a level equivalent to that in the spatial condition.

Understanding the precise nature of the coding underlying performance in the spatial and verbal conditions is a challenge. The analysis of initial errors involving only a reversal of direction along a given dimension gives some insight into this coding process. In Experiment 1 and in both sessions of Experiment 2, the right and left directions were confused only in the verbal condition, never in the spatial condition. Thus, there is some support for the earlier argument made by Farrell (1979) that understanding the words 'left' and 'right' requires the ability to 'tell' left from right, whereas copying movements does not. Also, the additional *post hoc* analyses revealed that these initial right/left direction confusions in the verbal condition did show the expected decline from the first half of the session to the second half of the session in both experiments.

Furthermore, the analysis of initial errors involving movement magnitude alone suggests that in the spatial condition, but not in the verbal condition, participants need to count the number of required movements, which imposes an added cognitive load when there are two movements. In contrast, the analysis of initial errors involving movement dimension alone revealed the same pattern for the verbal and spatial conditions. In both cases, participants confused the up/down dimension with the back/forward dimension, showing a preference to make vertical movements within one matrix rather than across matrices. Thus, although there are differences between the encoding of verbal and spatial information, there are also similarities. Further insights into the coding used in the verbal and spatial conditions might come from an individual difference analysis using measures of preferences for verbal and spatial strategies (see, e.g. Bacon, Handley, Dennis, & Newstead, 2008).

In any event, the present study has practical implications, both for the specific situation involving communication between pilots and air traffic control and, more generally, for the training, retention and transfer of the tasks of receiving and following navigation instructions, such as directions to rooms in a multi-story building or for ground transportation using Global Position Systems (GPSs). However, before such implications are reviewed several caveats need to be considered.

Although this task is meant to mimic the situation when a pilot receives navigation instructions from an air traffic controller, it differs from that situation in some respects. First, discrete movements are made in a two-dimensional grid rather than continuously in three-dimensional space, so the task can be viewed as requiring memory for a sequence of discrete locations rather than memory for continuous movements. Second, this paradigm does not use the specific nomenclature used for flight navigation and includes directions that are irrelevant for flight situations. For example, air traffic controllers never tell pilots to go backwards. However, a study using real pilots receiving realistic flight instructions in a flight simulator yielded results comparable to those found with college students in the present task (Mauro & Barshi, 2000). Also, Morrow, Lee, and Rodvold (1993) found that readback errors for controller-pilot communication concerning approach sectors were most frequent for runway identifications involving either the runway number or the left/right designation. Further, Loftus, Dark, and Williams (1979) have made the persuasive argument, 'It seems unlikely that human beings undergo fundamental changes in their ways of processing information by virtue of their being trained as pilots' (p. 180).

With these caveats in mind, the first practical implication is that whenever possible, navigation instructions should be limited to no more than three commands because of the steep decline in performance as a function of message length (see also, e.g. Cardosi, Brett, & Han, 1996, and Helleberg & Wickens, 2003, for similar observations based on analyses of readback errors in natural air traffic control communication and communication in an aviation simulator, respectively). The second implication is that optimal performance with

verbal navigation instructions requires practice translating the words heard to specified movements because of the striking improvement from the first to the second half of the session. The third implication, following from the observation that performance was better when the two conditions were the same than when they were switched across sessions, is that to maximize performance at test or in the field training should be conducted using the same type of instructions and displays as available during eventual application. The fourth implication applies to the case when the type of instructions and displays available at test or at eventual application are unknown. Because performance at test (averaged across test condition) was initially better following verbal training than following spatial training, training should be conducted using the more difficult verbal instructions under these circumstances, in agreement with earlier recommendations for training with 'desirable' difficulties (Bjork, 1994; McDaniel & Einstein, 2005; Schneider et al., 2002). Another possibility might be to combine verbal and spatial information during training (see Schneider, Healy, Buck-Gengler, Barshi, & Bourne, 2008), although no research has yet examined whether such a combination at training would be helpful in terms of transfer to a single modality at test.

ACKNOWLEDGEMENTS

This research was supported in part by Grants NCC2-1310, NNA05CS42A and NNA07CN59A from the National Aeronautics and Space Administration, Contract DASW01-03-K-0002 from the Army Research Institute and Grant W911NF-05-1-0153 from the Army Research Office to the University of Colorado. Portions of Experiment 1 were summarized at the 2004 convention of the American Psychological Association, and portions of Experiment 2 were summarized at the 2006 meeting of the Psychonomic Society. We thank Ernie Mross for help with data tabulation and Lyle Bourne for helpful comments about this research.

REFERENCES

Bacon, A. M., Handley, S. J., Dennis, I., & Newstead, S. E. (2008). Reasoning strategies: The role of working memory and verbal-spatial ability. *European Journal of Cognitive Psychology*, *20*, 1065–1086.

Baddeley, A. (2007). *Working memory, thought, and action*. New York: Oxford University Press.

Barshi, I., & Healy, A. F. (1998). Misunderstandings in voice communication: Effects of fluency in a second language. In A. F. Healy, & L. E. Bourne, Jr (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 161–192). Mahwah, NJ: Erlbaum.

Barshi, I., & Healy, A. F. (2002). The effects of mental representation on performance in a navigation task. *Memory & Cognition*, *30*, 1189–1203.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe, & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.

Brunyé, T. T., & Taylor, H. A. (2008a). Extended experience benefits spatial mental model development with route but not survey descriptions. *Acta Psychologica*, *127*, 340–354.

Brunyé, T. T., & Taylor, H. A. (2008b). Working memory in developing and applying mental models from spatial descriptions. *Journal of Memory and Language*, *58*, 701–729.

Cardosi, K. M., Brett, B., & Han, S. (1996). *An analysis of TRACON (terminal radar approach control) controller-pilot voice communications*. Cambridge: John A. Volpe National Transportation Systems Center (NTIS No. PB96-202593/HDM).

Corballis, M. C., & Beale, I. L. (1970). Bilateral symmetry and behavior. *Psychological Review*, *77*, 451–464.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–185.

Farrell, W. S. (1979). Coding left and right. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 42–51.

Healy, A. F. (2007). Transfer: Specificity and generality. In H. L. Roediger, III, Y. Dudai, & S. M. Fitzpatrick (Eds.), *Science of memory: Concepts* (pp. 271–275). New York: Oxford University Press.

Healy, A. F., & Nairne, J. S. (1985). Short-term memory processes in counting. *Cognitive Psychology*, *17*, 417–444.

Healy, A. F., Schneider, V. I., & Barshi, I. (2009). Cognitive processes in communication between pilots and air traffic control. In E. B. Hartonek (Ed.), *Experimental psychology research trends* (pp. 45–77). Hauppauge, NY: Nova Science Publishers.

Healy, A. F., Wohldmann, E. L., & Bourne, L. E., Jr. (2005). The procedural reinstatement principle: Studies on training, retention, and transfer. In A. F. Healy (Ed.), *Experimental cognitive psychology and its applications* (pp. 59–71). Washington, DC: American Psychological Association.

Healy, A. F., Wohldmann, E. L., Sutton, E. M., & Bourne, L. E., Jr. (2006). Specificity effects in training and transfer of speeded responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 534–546.

Helleberg, J. R., & Wickens, C. D. (2003). Effects of data-link modality and display redundancy on pilot performance: An attentional perspective. *International Journal of Aviation Psychology*, *13*, 189–210.

Loftus, G. R. (1978). Comprehending compass directions. *Memory & Cognition*, *6*, 416–422.

Loftus, G. R., Dark, V. J., & Williams, D. (1979). Short-term memory factors in ground controller/pilot communication. *Human Factors*, *21*, 169–181.

Logie, R. H., & Baddeley, A. D. (1987). Cognitive processes in counting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 310–326.

Maki, R. H. (1979). Right-left and up-down are equally discriminable in the absence of directional words. *Bulletin of the Psychonomic Society*, *14*, 181–184.

Mauro, R., & Barshi, I. (2000). Fear and flying: Effects of emotional stress on memory in simulated flight and analog tasks. In R. S. Jensen (Ed.), *Proceedings of the tenth symposium of aviation psychology* (pp. 1331–1336). Columbus, OH: The Ohio State University.

McDaniel, M. A., & Einstein, G. O. (2005). Material appropriate difficulty: A framework for determining when difficulty is desirable for improving learning. In A. F. Healy (Ed.), *Experimental cognitive psychology and its applications* (pp. 73–85). Washington, DC: American Psychological Association.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 519–533.

Morrow, D., Lee, A., & Rodvold, M. (1993). Analysis of problems in routine controller-pilot communication. *The International Journal of Aviation Psychology*, *3*, 285–302.

Mou, W., Zhang, K., & McNamara, T. P. (2004). Frames of reference in spatial memories acquired from language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 171–180.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*, 255–287.

Paivio, A. (2007). *Mind and its evolution: A dual-coding theoretical approach*. Mahwah, NJ: Erlbaum.

Schneider, V. I., Healy, A. F., & Barshi, I. (2004). Effects of instruction modality and readback on accuracy in following navigation commands. *Journal of Experimental Psychology: Applied*, *10*, 245–257.

Schneider, V. I., Healy, A. F., & Bourne, L. E., Jr. (2002). What is learned under difficult conditions is hard to forget: Contextual interference effects in foreign vocabulary acquisition, retention, and transfer. *Journal of Memory and Language*, *46*, 419–440.

Schneider, V. I., Healy, A. F., Buck-Gengler, C. J., Barshi, I., & Bourne, L. E., Jr. (2008). Effects of presenting navigation instructions twice in the same or different modalities. *Paper presented at the 49th Annual Meeting of the Psychonomic Society*, Chicago, IL, November.
Schneider, V. I., Healy, A. F., Kole, J. A., & Barshi, I. (2003). The verbal representation of navigation instructions depends on the spatial representation. *Invited poster presented at the 4th Tsukuba International Conference on Memory*, Tsukuba, Japan, January.
Sholl, M. J., & Egeth, H. E. (1981). Right-left confusion in the adult: A verbal labeling effect. *Memory & Cognition*, *9*, 339–350.
Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 352–373.
Wickens, C. D., Vidulich, M., & Sandry-Garza, D. (1984). Principles of S-C-R compatibility with spatial and verbal tasks: The role of display-control location and voice-interactive display-control interfacing. *Human Factors*, *26*, 533–543.

DOI: 10.1002/acp